

IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF WEST VIRGINIA

CHARLESTON DIVISION

LEE ANN SOMMERVILLE, et al.,

Plaintiffs,

v.

CIVIL ACTION NO. 2:19-cv-00878

UNION CARBIDE CORPORATION,

Defendant.

MEMORANDUM OPINION AND ORDER

Pending before the court are two motions to limit the opinions and testimony of Dr. Ranajit (Ron) Sahu filed by Defendants Union Carbide Corporation (“UCC”), [ECF No. 309], and Covestro LLC (“Covestro”), [ECF No. 327], respectively. Plaintiff Lee Ann Sommerville (“Plaintiff”) filed a response to UCC’s motion, [ECF No. 343], and a response to Covestro’s motion, [ECF No. 342], on July 10, 2023. For the reasons discussed below, I **GRANT** both motions, [ECF Nos. 309, 327], and **ORDER** that the testimony and reports of Dr. Ranajit Sahu be excluded as unreliable.

Upon detailed review of the expert reports, I find that the opinions of Dr. Sahu are not based upon sufficient facts or data because the inputs he uses in the air model are speculative and are premised on assumptions that do not accurately represent the Defendants’ operations in South Charleston. As such, there are serious doubts about whether such opinions are the product of reliable principles and methods,

which have also not been reliably applied to the facts of this case. For these reasons, which I discuss in greater detail below, I **GRANT** both motions, [ECF Nos. 309, 327], and **ORDER** that the testimony and reports of Dr. Ranajit Sahu be excluded as unreliable.

I. Background

The named Plaintiff brought this proposed medical monitoring class action against Defendants UCC and Covestro, as owners and operators of a manufacturing facility located in South Charleston, West Virginia, for their alleged “dangerous and reckless emission of ethylene oxide (“EtO”),” a known human carcinogen, between 1984 and 2019. [ECF No. 85, ¶¶ 1–2].

A. Ethylene Oxide Exposure

EtO is a colorless, odorless gas produced in large volumes at some chemical manufacturing facilities. *See Our Current Understanding of Ethylene Oxide (EtO)*, U.S. Env’t Prot. Agency, <https://www.epa.gov/hazardous-air-pollutants-ethylene-oxide/our-current-understanding-ethylene-oxide-eto#what> (last visited Mar. 11, 2024). According to the Environmental Protection Agency (“EPA”), EtO is a known human carcinogen, and regular, long-term exposure to EtO can cause certain cancers of the white blood cells, including non-Hodgkin lymphoma, myeloma, and lymphocytic leukemia. *Id.* Because the primary way EtO enters the environment is by release into the air, individuals who live near facilities that release EtO into the outdoor air may be exposed to EtO. *Id.*

According to Plaintiff, UCC is “the world’s leading producer of EtO,” and “the

South Charleston Plant is one of the only facilities in the [United States] that manufactures EtO.” [ECF No. 85, ¶ 28]. UCC has “owned and operated the South Charleston Plant since at least 1978,” and in 2015, Covestro began “operat[ing] polyols facilities” at the South Charleston plant. *Id.* ¶¶ 41, 44. Plaintiff claims that UCC and Covestro “operated without sufficient pollution controls to limit and/or eliminate the emissions of toxic EtO and, as a result, exposed thousands of residents in neighboring areas” to the carcinogen “for at least 41 years.” *Id.* ¶¶ 48, 53. As a result of the allegedly “negligent and tortious” emissions of EtO, Plaintiff claims that she and the proposed class members “have suffered significant exposure to hazardous EtO gases relative to the general population in the [United States],” *id.* ¶ 59, and are, therefore, “at an increased risk of developing cancer,” making “periodic diagnostic medical examinations [] reasonably necessary,” *id.* ¶ 62. Plaintiff seeks an award of “the quantifiable costs” of medical monitoring for the benefit of all proposed class members due to their alleged exposure. *Id.* ¶¶ 84–86.

B. Expert Opinion of Dr. Ranajit Sahu

Plaintiff relies upon a number of expert opinions in seeking to establish the elements of her claim. One such opinion is offered by Dr. Ranajit Sahu, a mechanical engineer, who was engaged to provide “technical expertise, analysis, methodology, and opinions regarding various environmental and pollutant fate and transport issues relating to emissions of [EtO] from Union Carbide’s Institute and [the] South Charleston facilities” (collectively, the “Facilities”) to ultimately determine the potential exposure levels of Plaintiff and the proposed class members. [ECF No. 309-

3, at 6]. Dr. Sahu provided his Expert Report on February 17, 2023. [ECF No. 309-3]. Defendants seek to exclude the opinions in this report, as well as Dr. Sahu's related testimony.

1. Dr. Sahu's Background and Experience

Dr. Sahu is "a mechanical engineer with over thirty years of experience in the fields of environmental, mechanical, and chemical engineering." *Id.* After obtaining a Bachelor of Technology in mechanical engineering from the Indian Institute of Technology in Kharagpur, India in 1983, Dr. Sahu attended the California Institute of Technology in Pasadena, California from 1984 to 1988, where he received a Master of Science and Doctorate degree in mechanical engineering. *Id.* at 7. Dr. Sahu also has over twenty-seven years of "project management experience" and has "provided consulting services to numerous private sector, public sector, and public interest group clients" over the past twenty-five years. *Id.* His experiences include "design and specification of pollution control equipment for a wide range of emissions sources," "multimedia environmental regulatory compliance," "transportation air quality impact analysis," "air quality . . . permitting," "multimedia/multi-pathway human health risk assessment for toxics," and "air dispersion modeling." *Id.* at 6–7. Dr. Sahu has also taught numerous university courses in air pollution, process hazard analysis, air pollution controls, transportation and air quality, and engineering. *Id.* at 7.

2. Dr. Sahu's Summary of Opinions

In his February 17, 2023, report, Dr. Sahu outlines his findings related to the EtO emissions from operations at the Facilities. His summary of opinions is listed

below:

(i) There were numerous sources of EtO air emissions from Union Carbide's West Virginia Operations (WVO). EtO was used extensively in various production processes. EtO was brought into the [Facilities] from elsewhere, and subsequently distributed, stored, and processed. This handling and use of EtO created significant opportunities for EtO releases into the ambient air. Examples include: EtO emissions from various storage tanks, where EtO was present in the materials being stored, and therefore emitted from such tanks; EtO emissions from various fugitive components—such as valves, pump seals, connectors, flanges, and others—which were present throughout the [Facilities], including in all handling, storage, distribution, and processing areas; EtO emissions from so-called extended cookout operations at the reactors; and EtO emissions from flaring of waste gases containing EtO.

(ii) In spite of the many sources of EtO at the [Facilities], there was no direct testing of EtO mass emissions to the ambient air from any of the sources under representative conditions over the many years of EtO-related operations.

While direct monitoring of the open-flame stack flares at the [Facilities] would have been problematic, no efforts were made to use other types of thermal combustion devices (such as enclosed stack flares, thermal oxidizers, etc.) for waste gas disposal, which could have been tested and monitored.

There was no direct testing of the many fugitive components that emitted EtO. The Leak Detection and Repair (LDAR) program that was used for some period of time relied on assumed emission factors and effectiveness of the program with no verification.

There was no testing of EtO emissions from any of the tanks that stored EtO containing materials.

There was little direct testing of EtO from extended cookout operations, which were claimed to be appropriate regulatory controls for various hazardous air pollutants, including EtO.

(iii) Since there was no direct testing of EtO emissions from the many sources of EtO, and certainly therefore no testing under representative conditions, all emissions of EtO developed and/or reported by Defendants used engineering estimates. However, for engineering estimates to be proper, they needed to be grounded by site specific data.

Site specific data is not constant at chemical manufacturing operations such as at the operations in question. Union Carbide's estimates of EtO emissions relied on sporadic and unsupported assumptions; and, crucially, in the case of flares, on "surrogate" parameters that bore no relation to actual operating conditions or actual emissions.

As such it is my opinion that EtO emissions were underestimated significantly from all sources. In this report, I have highlighted the example of underestimation of the EtO emissions from the flares, via the use of an unrealistically high flare control efficiency, which did not and could not have represented every-day operations of the flares.

(iv) The few verification tests that were conducted at the facilities in order to meet regulatory requirements showed that actual operations deviated considerably from regulatory requirements. In the case of the flares, preliminary tests conducted prior to the required compliance tests, showed significant failures. Due to the lack of prior testing at the facilities, these significant failures indicate that inadequate conditions existed for persistent periods of time prior to the failed tests. This is supported by internal discussions among Union Carbide managers explicitly acknowledging that the control efficiencies they reported to regulators rarely matched reality.

(v) Failed tests were remedied for regulatory compliance demonstrations in order to make them "pass." However, the failed tests were never reported to regulators as required. Once a "passing" test was obtained, there is no record that Union Carbide monitored subsequent flare operating conditions, or had any policies or procedures to ensure that necessary passing conditions were maintained. Thus, the compliance tests were unrepresentative of usual or normal operations and were meaningless. Worse, it is my opinion, that by conducting sham compliance tests, Defendants were in violation of applicable regulatory requirements, such as the need to conduct testing under representative conditions and to keep proper records demonstrating continuous compliance.

(vi) Based on conservative (i.e., still likely underestimated emissions) recalculation of emissions and by modeling such emissions using appropriate and standard air dispersion modeling, it is my opinion that persons in the surrounding community were exposed to significant concentrations of EtO.

(vii) Collectively, it is my opinion that Defendants' conduct relating to the emissions of EtO was irresponsible and exhibited a knowing and

willful disregard of both their regulatory requirements and the health and safety of persons in the surrounding community.

Id. at 8–10.

C. Defendant UCC's Motion to Exclude [ECF No. 309]

UCC asserts that Dr. Sahu should be excluded as a witness for four primary reasons. First, UCC argues that Dr. Sahu's air modeling opinions do not "fit" the facts of the case. [ECF No. 310, at 3–4].¹ Second, UCC asserts that Dr. Sahu's air modeling opinions suffer from "serious methodological flaws" which render them unreliable because they do not rely on an accurate historical reconstruction and were formulated using "patently unreliable emissions data" that Dr. Sahu failed to validate. *Id.* at 11, 13, 15. Third, UCC contends that Dr. Sahu's source configurations do not reliably represent historical conditions. *Id.* at 19. Finally, UCC claims that Dr. Sahu did not employ a reliable method to estimate background levels of EtO in the community and combined contradictory meteorological data that leads to impossible results. *Id.* at 20. For these reasons, UCC argues that "Dr. Sahu's work . . . falls far short of the standard set by this [c]ourt," *id.* at 3, and should be excluded pursuant to Rule 702 of the Federal Rules of Evidence and *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993).

Plaintiff rebuts these arguments, arguing that UCC "does not mount a true *Daubert* challenge." [ECF No. 343, at 3 (quoting *TWFS, Inc. v. Shaefer*, 325 F.3d 234, 240 (4th Cir. 2003))]. Rather, Plaintiff claims that UCC "mounts an attack on the

¹ Unless otherwise stated, the page numbers cited correspond to CM/ECF page numbers.

data inputs to Dr. Sahu's air model by disputing the accuracy of the data and the assumptions that are required to fill gaps in the data, while failing to offer any different conclusions that might be reached with different data." *Id.* She argues that the court should deny UCC's motion because "questions regarding the factual underpinnings of an expert's opinion affect the weight and credibility of the opinion, not admissibility." *Id.* (citing *Bresler v. Wilmington Tr. Co.*, 855 F.3d 178, 195 (4th Cir. 2017)).

D. Defendant Covestro's Motion to Exclude [ECF No. 327]

Likewise, Defendant Covestro asserts that the court should exclude the testimony and reports of Dr. Sahu because they are irrelevant to Plaintiff's claims against Covestro and are also fatally unreliable. [ECF No. 327, at 1]. Specifically, Covestro contends that (1) the opinions are "irrelevant as to Covestro" because Dr. Sahu is "unfamiliar" with Covestro's operations in South Charleston; (2) the air model is "irrelevant and unreliable as to Covestro because its inputs do not reflect the reality" of its operations in South Charleston; (3) the figures produced by Dr. Sahu's air model are "irrelevant and unreliable because they do not represent exposure or dose;" and (4) the opinions "regarding the exposure of the general population" are "irrelevant and unreliable because Dr. Sahu does not know the background levels of [EtO] in the air." *Id.*

In response, Plaintiff argues that Covestro, too, fails to "mount a true *Daubert* challenge," but instead, "challenges the factual underpinnings, data inputs, and relevancy of Dr. Sahu's opinions." [ECF No. 342, at 3]. She claims that Dr. Sahu had

properly considered Covestro’s South Charleston operations and that his opinions are based on those operations. *Id.* She also argues that Dr. Sahu’s opinions on “significant EtO exposure are relevant,” that the “meteorological inputs are reliable,” and that the EPA’s National Air Toxics Assessment (“NATA”) database—which Dr. Sahu used as a data source—“provides the most appropriate background level.” *Id.* Further, while Plaintiff concedes that there was a “minor error with a data input” related to Covestro’s operations, she asserts that “Dr. Sahu properly updated his expert report” to correct the issue,² and that regardless, the “minor error” was not sufficient grounds to exclude his opinions. *Id.* For these reasons, Plaintiff asks the court to deny Covestro’s motion. *Id.* at 12.

II. Legal Standard

Under Federal Rule of Evidence 702, expert testimony is admissible if it (1) will “help the trier of fact to understand the evidence or to determine a fact in issue” (2) is “based upon sufficient facts or data,” and (3) is “the product of reliable principles and methods,” which (4) have been “reliably appli[ed] . . . to the facts of the case.” Fed. R. Evid. 702. A two-part test governs the admissibility of expert testimony: evidence is admissible if “it rests on a reliable foundation and is relevant.” *Daubert*, 509 U.S. at 597 (1993). The proponent of expert testimony does not have the burden to “prove” anything. He must, however, “come forward with evidence from which the court can determine that the proffered testimony is properly admissible.” *Md. Cas.*

² There is currently pending a motion filed by Covestro to strike this supplemental expert report as untimely under Federal Rule of Civil Procedure 26(e). *See* [ECF No. 300].

Co. v. Therm-O-Disc., Inc., 137 F.3d 780, 783 (4th Cir. 1998).

The district court has an important role as the gatekeeper of expert opinions and testimony. Because “expert witnesses have the potential to be both powerful and quite misleading,” the court must “ensure that any and all scientific testimony . . . is not only relevant, but reliable.” *Cooper v. Smith & Nephew, Inc.*, 259 F.3d 194, 199 (4th Cir. 2001) (internal markings omitted) (citing *Westberry v. Gislaved Gummi AB*, 178 F.3d 257, 261 (4th Cir. 1999) and *Daubert*, 509 U.S. at 588, 595). I “need not determine that the proffered expert testimony is irrefutable or certainly correct”— “[a]s with all other admissible evidence, expert testimony is subject to testing by ‘[v]igorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof.’” *United States v. Moreland*, 437 F.3d 424, 431 (4th Cir. 2006) (quoting *Daubert*, 509 U.S. at 596) (alteration in original); *see also Md. Cas. Co.*, 137 F.3d at 783 (noting that “[a]ll *Daubert* demands is that the trial judge make a ‘preliminary assessment’ of whether the proffered testimony is both reliable . . . and helpful”).

Daubert outlines specific factors to guide the overall relevance and reliability determinations that apply to all expert evidence. They include “(1) whether the particular scientific theory can ‘be (and has been) tested’; (2) whether the theory ‘has been subjected to peer review and publication’; (3) the ‘known or potential rate of error’; (4) the ‘existence and maintenance of standards controlling the technique’s operation’; and (5) whether the technique has achieved ‘general acceptance’ in the relevant scientific or expert community.” *United States v. Crisp*, 324 F.3d 261, 266

(4th Cir. 2003) (quoting *Daubert*, 509 U.S. at 593–94).

With respect to relevancy, *Daubert* explains:

Expert testimony which does not relate to any issue in the case is not relevant and, ergo, non-helpful. The consideration has been aptly described by Judge Becker as one of “fit.” “Fit” is not always obvious, and scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes. . . . Rule 702’s “helpfulness” standard requires a valid scientific connection to the pertinent inquiry as a precondition to admissibility.

Id. at 591–92 (internal citations omitted).

Despite these factors, “[t]he inquiry to be undertaken by the district court is ‘a flexible one’ focusing on the ‘principles and methodology’ employed by the expert, not on the conclusions reached.” *Westberry*, 178 F.3d at 261 (quoting *Daubert*, 509 U.S. at 594–95); *see also Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 150 (1999) (“We agree with the Solicitor General that ‘[t]he factors identified in *Daubert* may or may not be pertinent in assessing reliability, depending on the nature of the issue, the expert’s particular expertise, and the subject of his testimony.’”) (citation omitted); *see also Crisp*, 324 F.3d at 266 (citing *Kumho* for the proposition “that testing of reliability should be flexible and that *Daubert*’s five factors neither necessarily nor exclusively apply to every expert”). However, “conclusions and methodology are not entirely distinct from one another.” *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997).

Additionally, some scholars have opined that “[*Daubert*] has proven especially inappropriate and inadequate in evaluating model reliability. The problems that plague all models—namely uncertainty and a lack of transparency—are exacerbated in environmental-tort litigation.” Matthew W. Swinehart, *Remedying Daubert’s*

Inadequacy in Evaluating the Admissibility of Scientific Models Used in Environmental-Tort Litigation, 86 Tex. L. Rev. 1281, 1283 (2008) [hereinafter *Remedying Daubert's Inadequacy*]; see also Wendy Wagner et al., *Misunderstanding Models in Environmental and Public Health Regulation*, 18 N.Y.U. Env't'l L.J. 293, 324 (2010) [hereinafter *Misunderstanding Models*] (“There are significant conflicts between the Supreme Court’s deterministic test for admissibility of expert testimony (embodied in *Daubert v. Merrell Dow Pharmaceuticals*) and the true nature of models.”). Ultimately, a district court has broad discretion in determining whether to admit or exclude expert testimony, and “the trial judge must have considerable leeway in deciding in a particular case how to go about determining whether particular expert testimony is reliable.” *Cooper*, 259 F.3d at 200 (quoting *Kumho Tire*, 526 U.S. at 152); see also *United States v. Hankey*, 203 F.3d 1160, 1167 (9th Cir. 2000) (“A trial court not only has broad latitude in determining whether an expert’s testimony is reliable, but also in deciding *how* to determine the testimony’s reliability.”) (internal quotation marks and citations omitted).

III. Discussion

Before reaching the analysis of Dr. Sahu’s expert opinions, I will first address Plaintiff’s contention that the instant motions raise questions affecting *the weight and credibility* of Dr. Sahu’s opinions rather than their *admissibility* and the inference that the issues are, thus, a matter for the jury, not the court. See [ECF No. 343, at 3].

As part of its “gatekeeping function,” the court must “ensure[] that expert

evidence is sufficiently relevant and reliable when it is submitted to the jury.” *Nease v. Ford Motor Co.*, 848 F.3d 219, 231 (4th Cir. 2017) (emphasis in original). “Thus, the trial court must probe the reliability and relevance of expert testimony any time ‘such testimony’s factual basis, data, principles, methods, or their application are sufficiently called into question.” *EEOC v. Freeman*, 778 F.3d 463, 472 (4th Cir. 2015) (Agee, J., concurring) (quoting *Kumho*, 526 U.S. at 149). “[T]he fact that an expert witness [is] ‘subject to a thorough and extensive examination’ does not ensure the *reliability* of the expert’s testimony; such testimony must still be assessed *before* it is presented to the jury.” *Nease*, 848 F.3d at 231 (quoting *McClain v. Metabolife Int’l, Inc.*, 401 F.3d 1233, 1238 (11th Cir. 2005) (emphasis added)). Therefore, “[a]lthough the jury ultimately decides the ‘weight’ of the evidence, the judge ensures there is sufficient probative value . . . to justify submitting the issue in the first instance.” Daniel D. Blinka, *Expert Testimony and the Relevancy Rule in the Age of Daubert*, 90 Marq. L. Rev. 173, 191 (2006). Here, the issues presented to the court deal not only with the validity of the underlying data used by Dr. Sahu but also with Dr. Sahu’s methods in applying that data to the facts of the case. I will now proceed to Defendants’ arguments regarding relevance and reliability.

A. Relevance Challenge: The “Fit” of Dr. Sahu’s Opinions to the Facts of the Case

First, both UCC and Covestro argue that Dr. Sahu’s opinions are irrelevant or do not “fit” the facts of the case, because this case deals solely with EtO emissions from the Defendants’ operations at South Charleston, and Dr. Sahu’s model consists of combined emissions sources from operations at both Institute *and* South

Charleston. *See* [ECF No. 310, at 12]; [ECF No. 328, at 4]. According to UCC, this failure to disaggregate the emissions sources and resulting exposure estimates makes it “impossible to discern the impacts of EtO emissions” from just the South Charleston facility. [ECF No. 310, at 3]. And Defendants argue that this same issue plagues the jury’s ability to discern the alleged impacts from each of their separate operations at that facility. As a result, Defendants claim that Dr. Sahu’s model and opinions are irrelevant and should be excluded. *Id.* at 12; *see* [ECF No. 328 at 13 (“Dr. Sahu’s opinions are irrelevant and unreliable because he does not . . . separate alleged impacts between the defendants.”)].

Plaintiff, conversely, argues that while Dr. Sahu’s model presents combined results from both Institute and South Charleston, it is “entirely capable of predicting, at each desired receptor, the impacts from not just one plant but from specific sources within each plant.” [ECF No. 342, at 9 (quoting [ECF No. 327-13, at 17])]. Further, Plaintiff contends that Dr. Sahu’s opinions are relevant because his model shows that proposed class members “were ‘significantly exposed’ to a hazardous substance relative to the general population.” *Id.* (quoting *In re W. Va. Rezulin Litig.*, 585 S.E.2d 52, 73 (W. Va. 2003)). These facts, Plaintiff argues, defeat Defendants’ relevancy challenges.³

³ Plaintiff also argues that because West Virginia law allows plaintiffs to sue “only one of several joint tortfeasors,” *Bradley v. Appalachian Power Co.*, 256 S.E.2d 879, 886 (W. Va. 1979), and does not require plaintiffs to “prove that the negligence of one of the defendants is the sole proximate cause of the injury,” *Kodym v. Frazier*, 412 S.E.2d 219, 224 (W. Va. 1991), the Defendants here cannot avoid liability by “blaming the emissions from its other plant.” [ECF No. 343, at 6]. However, because that argument deals with liability, rather than the *relevance* of Dr. Sahu’s report, I need not address it here.

Under *Daubert*, expert testimony should be “sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute.” *Daubert*, 509 U.S. at 591 (quoting *United States v. Downing*, 753 F.2d 1224, 1242 (3d Cir. 1985)); *see also* *United States v. Chapman*, 209 F. App’x 253, 269 (4th Cir. 2006) (noting that expert testimony is “excludable under Rule 702 if it does not aid the jury”). While I agree that Dr. Sahu’s failure to disaggregate the sources of EtO emissions in his report creates issues of reliability, I cannot agree that this renders the report or Dr. Sahu’s opinions entirely irrelevant. Dr. Sahu’s opinions are inarguably related to the issues in this case even if the report contains superfluous information. Therefore, I find that the opinions of Dr. Sahu are relevant and decline to exclude his opinions on this ground.

B. Reliability Challenges: Dr. Sahu’s “Serious Methodological Flaws”

Next, Defendants UCC and Covestro argue that Dr. Sahu’s opinions suffer from “serious methodological flaws” and are, therefore, unreliable. [ECF Nos. 310, at 4]; *see also* [ECF No. 328, at 3–4]. I agree. The issues plaguing Dr. Sahu’s opinions stem not only from concerns about the validity of the underlying data but also with the methodology used by Dr. Sahu’s in *applying* that data.

“Evidence is admissible only if ‘it rests on a reliable foundation.’” *Freeman*, 778 F.3d at 471 (quoting *Daubert*, 509 U.S. at 597). Courts must, therefore, verify that expert testimony is “based on sufficient facts or data.” *See* Fed. R. Evid. 702(b). And courts widely agree that “trial judges may evaluate the data offered to support an expert's bottom-line opinions to determine if that data provides adequate support to

mark the expert's testimony as reliable.” *Milward v. Acuity Specialty Prods. Grp., Inc.*, 639 F.3d 11, 15 (1st Cir. 2011) (cleaned up); *accord Blunt v. Lower Merion Sch. Dist.*, 767 F.3d 247, 276 (3d Cir. 2014); *In re TMI Litig.*, 193 F.3d 613, 697 (3d Cir. 1999); *United States v. City of Miami*, 115 F.3d 870, 873 (11th Cir. 1997).

“Dispersion modeling is considered to be both an art and a science. There are numerous variables to consider in any such analysis, each of which may be open to careful scrutiny and analysis during the litigation process.” Andrew C. Wilson et al., *Tracking Spills and Releases: High-Tech in the Courtroom*, 10 Tul. Env’t L.J. 371, 390 (1997) [hereinafter Wilson et al., *Tracking Spills and Releases*].

While the “troubled marriage between science and law” is now a familiar phenomenon, models may best be thought of as the unhappy couple’s inconsolable child. All of the challenges that afflict science used for policy also afflict models: significant uncertainties and subjective judgments loom in model design, in interpretation of the data, and in communicating that data to other scientists as well as lay persons.

Wagner et al., *Misunderstanding Models*, *supra*, at 308; *see also* Swinehart, *Remedying Daubert’s Inadequacy*, *supra*, at 1288–89 (“Model development is plagued by uncertainties regarding (A) which components of a system to include in the model; [and] (B) whether one has gathered accurate, precise, and representative data about the system . . .”) (internal quotation marks and citations omitted). Additionally,

. . . models used in environmental-tort litigation present two additional issues. First, the adversarial process creates unfortunate incentives to misuse models in “unscientific” ways, perhaps taking advantage of their uncertain and impervious nature. Second, because models are particularly complex and enigmatic applications of science, they are even less accessible to judges

Swinehart, *Remedying Daubert’s Inadequacy*, *supra*, at 1288. Air models, by their

nature, have “many moving parts.” *Coleman v. Union Carbide Corp.*, No. 2:11-0366, 2013 WL 5461855, at *23 (S.D. W. Va. Sept. 30, 2013). “Modeling involves synthesizing disparate data, assumptions, uncertainties, and theories in the most robust way possible.” Wagner et al., *Misunderstanding Models*, *supra*, at 310. “The accuracy of the model bears a strong positive relationship to the correct inputs being used—inputs that represent the actual conditions at the facility and its emission sources.” *Coleman*, 2013 WL 5461855, at *23; *see also* Swinehart, *Remedying Daubert’s Inadequacy*, *supra*, at 1294–95 (“A model is only as good as its underlying data; even the best models cannot make up for faulty data.”); Charles D. Case, *Problems in Judicial Review Arising from the Use of Computer Models and Other Quantitative Methodologies in Environmental Decisionmaking*, 10 Bos. Coll. Env’t Aff. L. Rev. 251, 275 (1982) (“[A] lack of data can hamper the accurate application of a model.”); Thomas O. McGarity & Wendy E. Wagner, *Legal Aspects of the Regulatory Use of Environmental Modeling*, 33 ELR 10751, 10770 (2003) (“[T]he empirical data upon which models are based and the input data from which models extrapolate to make predictions are critical to the accuracy of those predictions.”).

UCC argues that Dr. Sahu’s air model does not accurately represent “the actual conditions” at the South Charleston plant or the “emission sources over the 36-year period” at issue in this case. [ECF No. 310, at 5]. Specifically, UCC takes issue with Dr. Sahu’s (1) use of “maximum possible emissions figures from permit applications” in lieu of actual emissions data; (2) “cherry-picking” of Institute plant flare emissions from one year and “counter-factually assuming that the emissions

from that sources remained constant over an 18-year period;” and (3) “making the ahistorical assumption that current-day operations existed unchanged over the 36-year period.” *Id.* Additionally, UCC contends that “Dr. Sahu use[s] three different sets of meteorological data from different locations with different wind directions and speeds, and then added the concentrations together to reach an impossible model result.” *Id.* Due to “the lack of reliability of his model inputs, including source configuration, emissions characterization, and meteorological conditions, and no attempt to validate his model results,” UCC argues that “Sahu’s modeling results are unreliable and not scientifically defensible.” [ECF No. 309-1, at 8].

Likewise, Covestro seeks to exclude Dr. Sahu’s opinions and testimony as unreliable. Specifically, Covestro contends that Dr. Sahu’s “constant emission rates *overstate[] ethylene oxide emissions by 30%* compared to Covestro’s reported emissions, which themselves are overestimates.” [ECF No. 328, at 4 (emphasis in original)]. Additionally, Covestro argues that “Dr. Sahu also model[s] emissions from Covestro sources going back to 1984, when Covestro did not begin its South Charleston operations until April 1, 2000.” *Id.* Finally, Covestro claims that Dr. Sahu’s background concentration level is unreliable and that his modeled results “do not represent exposure or dose.” *Id.* As such, Covestro asks the court to exclude Dr. Sahu’s opinions and testimony.

Because both Defendants’ motions deal with the same substantive issues concerning reliability, I will address them together.

1. Failure to Represent Actual Emissions at South Charleston

i. Inapplicable Emissions Sources

UCC employed its own expert, Ranjit J. Machado, P.E. (“Machado”), to examine Dr. Sahu’s report.⁴ According to Machado, Dr. Sahu models four groups of sources in this case: (1) UCC’s and Covestro’s South Charleston facility sources, not including tanks; (2) UCC storage tanks; (3) Institute facility sources, not including flares; and (4) Institute facility flares, “each with their own set of meteorology.” [ECF No. 309-1, at 38]. UCC claims that “Dr. Sahu’s results do not differentiate impacts from each of these source groups but instead only present the final combined concentrations, meaning it is impossible to discern the impacts from the individual source groups from Dr. Sahu’s presented results.” [ECF No. 310, at 9]. This is important, UCC alleges, because in this lawsuit, “Plaintiff has only alleged exposure to emissions from the South Charleston facility, not the Institute facility.” *Id.* For instance, Dr. Sahu uses flare emission data from the Institute facility, which “account[s] for the overwhelming majority of modeled air concentrations” and has a “significant impact” on Dr. Sahu’s exposure conclusions. *Id.* at 9, 12. But that emissions data does not apply to the South Charleston facility, as it has never used flares. *Id.* UCC argues that this methodology makes “it is impossible to draw conclusions about exposure from South Charleston alone from Dr. Sahu’s results since Institute disproportionately impacts concentrations across the domain.” *Id.* at

⁴ Although I discuss Machado’s report while explaining the challenges made to Dr. Sahu’s model, my decision to exclude the opinions of Dr. Sahu rests solely on the law and research detailed above as applied in this case.

13.

In Dr. Sahu’s expert rebuttal report, he argues that “[w]hile [he] has presented the results from both plants, the model is entirely capable of predicting, at each desired receptor, the impacts from not just one plant but from specific sources within each plant.” [ECF No. 309-6, at 20]. Thus, he claims, “there is no issue about predicting impacts from Institute, impacts from [UCC’s] operations at South Charleston[,] and Covestro’s impacts from its operations at South Charleston separately.” *Id.*

Dr. Sahu’s model and expert report, as provided to Defendants—and, in turn, the court—do not disaggregate the impacts between facilities or separate operations at the South Charleston facility. The court cannot, therefore, determine the reliability of the estimated exposure from the Defendants’ operations based on the information provided. Furthermore, the model and report suffer from additional methodological flaws that further merit exclusion.

ii. Assumptions in Historical Data

The next challenge to the Dr. Sahu’s methodology is that he assumes historical data—rather than using records that reflect actual conditions at the Facilities—and fails to verify that data.

Experts identify emissions sources “by reviewing plot plans, process diagrams, and operational information for the specific time period involved.” *Coleman*, 2013 WL 5461855, at *23. After collecting this data, “one then determines the amount of targeted pollutants produced from each source.” *Id.* The process requires “[a]n in-

depth data investigation, a searching historical analysis, an excruciating attention to detail, and a methodology designed to wring error out of the process” because with the model, “one is attempting to recreate decades of emissions, plume movements, and particle depositions.” *Id.* at *24. Dr. Sahu did not complete this process and instead makes assumptions regarding historical data at the Facilities with little to no basis for doing so.

(a) Source Parameters

First, in defining source parameters and configurations for the full period of 1984 to 2019, Dr. Sahu uses modeling conducted in 2019 by the West Virginia Department of Environmental Protection (“WVDEP”) and, according to UCC, “assume[s] that during that 36-year period, there were no relevant changes in those source characteristics or other plant operations.” [ECF No. 310, at 11]. For areas not covered by this WVDEP model, UCC contends that Dr. Sahu “invent[s] volume sources ‘covering the general area’ at the plants where that process area had previously operated.” *Id.* (citing [ECF No. 309-3, at 63 (noting that where the 2019 WVDEP modeling could not be used, “a volume source was used for modeling purposes covering the general area”)]); *see* [ECF No. 309-1, at 8 (arguing that Dr. Sahu “does not account for changes in source characteristics over time and artificially creates source parameters for historical sources that were not modeled by WVDEP, rather than providing factual model parameters as model inputs”)]. For example, for areas “where WVDEP/UCC prepared fugitive emission estimates,” UCC argues that Dr. Sahu “took an average” of those estimates “and applie[s] this same distribution”

to all of the source areas for “all years where . . . fugitive emissions were estimated.” [ECF No. 309-3, at 63].

Additionally, for the period from 1985 to 1989, Dr. Sahu models all South Charleston emissions as fugitive emissions, which UCC contends was incorrect because there is “no reason to believe that no point sources existed during that time frame.”⁵ [ECF No. 310, at 11 (citing [ECF No. 309-3, at 58]; [ECF No. 309-2, at 36–37 (admitting he modeled all of the sources for 1985 to 1988 as fugitive emission sources))]. UCC asserts that Dr. Sahu “assume[s] without support that emissions were apportioned between fugitive and point sources in the exact same proportion [for the years 1990 to 1995] as they were in 1996.” *Id.* (citing [ECF No. 309-3, at 62]). Finally, UCC’s expert, Machado, asserts that “[a]lgorithms in air dispersion models depend on accurate source parameter information to reliably simulate dispersion.” [ECF No. 309-1, at 8]. “These parameters, including temperature, exit velocity, stack/release height, and diameter, are used with the model’s pre-defined algorithms that calculate emission plume dispersion, and in turn, modeled concentrations.” *Id.* at 11. Machado suggests that it is crucial to apply a “rigorous methodology to accurately represent each emission source” when producing air dispersion modeling

⁵ Point source emissions, “also called ‘stack emissions,’” are defined as “[a]ir releases that occur through [identifiable] confined air streams, such as stacks, ducts or pipes,” whereas fugitive emissions are those “[r]eleases to air that don’t occur through a confined air stream” and includes “equipment leaks, releases from building ventilation systems and evaporative losses from surface impoundments and spills.” *Toxics Release Inventory (TRI) Program: Common TRI Terms*, U.S. Env’t Prot. Agency, <https://www.epa.gov/toxics-release-inventory-tri-program/common-tri-terms> (last visited Mar. 14, 2024).

predictions.⁶ *Id.*

Dr. Sahu disagrees with these sentiments, stating that “[f]or modeling purposes” certain sources “can be modeled in a variety of ways.” [ECF No. 309-6, at 13]. He claims that “[t]here is nothing extraordinary or ‘contrived’ about this.” *Id.* Instead, he blames the Defendants for his need for assumptions and any lack of precision in determining source parameters. *See id.* (“Such sources may have been capable of being modeled with greater precision had defendants taken efforts to describe these sources more precisely in the ordinary course of business—but they did not.”). He goes on to claim that “even if these sources were capable of being modeled more precisely, it is unlikely this would have materially changed the results of the analysis.” *Id.* When asked for the basis for his assertion that more precise source parameters would not have a material effect on the results of his analysis, Dr. Sahu states that it is simply “[his] judgment.” [ECF No. 309-2, at 27]. He goes on to argue that “the easiest way to prove [him] wrong would have been to show[] and correct all the criticisms with an alternate set of analyses.”⁷ *Id.* Further, he states that it is his “judgment” that “even if [the source parameters] were ‘corrected’ in some

⁶ Machado provides the following reasoning for his assertion:

The high temperature of flare stacks will cause emission plumes to have buoyancy, which gives way to “plume rise” that disperses emissions vertically. Conversely, fugitive emissions are often released at ambient temperature, leading to less buoyancy, less vertical dispersion, and higher modeled ground concentrations near the emission source. . . . Two sources with equal emissions rates, but with different source configurations, . . . can result in orders of magnitude differences in offsite concentrations.

[ECF No. 309-1, at 11].

⁷ Because it is the proffered expert’s duty to illustrate that his assumptions are “more than subjective belief or unsupported speculation,” I find Dr. Sahu’s argument that Defendants failed to prove him wrong or provide alternative analyses to be unconvincing. *See Daubert*, 509 U.S. at 590 (stating that “the word ‘knowledge’ connotes more than subjective belief or unsupported speculation”).

fashion with a whole bunch of unsupported *alternative assumptions*, they would not have changed the conclusions” that Dr. Sahu reached. *Id.* (emphasis added). He concedes, however, that despite Machado’s criticisms, he “ha[s] not done any additional analysis” or new modeling to rebut Machado’s assertions. *Id.*

“The accuracy of the model depends upon the rigor applied in the input gathering process.” *Coleman*, 2013 WL 5461855, at *24; *see also* Wilson et al., *Tracking Spills and Releases, supra*, at 390 (“The method of application of the data is as important to the modeling process as the integrity of the data being used.”). Although Dr. Sahu attempts to blame Defendants for the lack of reliable, site-specific data to base his model on, it is clear that he made assumptions in determining the source parameters for his model with little to no scientific basis for doing so. Simply put, the data Dr. Sahu used for the air model is based on nothing but Dr. Sahu’s own assertions, and the court will not presume such opinions to be reliable. *See Gen. Elec. Co.*, 522 U.S. at 146 (“[N]othing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert.”); *see also Coleman*, 2013 WL 5461855, at *31 (stating that “emission rates from actual facility records are far superior to hypothesized extrapolations”).

Despite Dr. Sahu’s contention that more precise source parameters would not have a material effect on the results of his analysis, this court has previously made clear that “the location and configuration of emissions sources if *of critical importance* in producing an accurate air dispersion model.” *Coleman*, 2013 WL 5461855, at *25

(emphasis added). Indeed, “it is critical in the air modeling process to *accurately* estimate (1) emissions, (2) their sources and the dispersion characteristics of the sources, and (3) the location, size, and physical characteristics of the sources.” *Id.* at *28 (emphasis added); *see also Gulf S. Insulation v. U.S. Consumer Prod. Safety Comm’n*, 701 F.2d 1137, 1146 (5th Cir. 1983) (“To make precise estimates, precise data [is] required.”). Dr. Sahu has not done that.

(b) Emissions Data by Period

Next, in addition to such assumptions, Defendants argue that Dr. Sahu also uses unreliable or unvalidated emissions data “for different years throughout the period he models” and makes several assumptions in applying that data to relevant periods. [ECF No. 310, at 10–11]. Covestro argues that “unburdened by information contained in Covestro’s timely produced documents or known by Covestro’s witnesses, Dr. Sahu model[s] Covestro’s [EtO] emissions by applying emission rates that he [holds] constant for 1984 to 2019.” [ECF No. 328, at 10].

“[I]t is methodologically unsound to assume [a manufacturing plant] operated historically in a static sense.” *Coleman*, 2013 WL 5461855, at *30. But that is precisely what Dr. Sahu does here, despite the fact that Covestro did not begin operating at the South Charleston facility until 2000. *See* [ECF No. 328, at 11]. In short, Dr. Sahu chooses “the far simpler, but inherently unreliable, approach of assuming” that operations at the Facilities remained “essentially unchanged historically” rather than examining their actual operations histories. *See Coleman*, 2013 WL 5461855, at *29. This, along with his unsubstantiated assumptions outlined

below, undermines the reliability of Dr. Sahu's findings.

a. 1984

According to Dr. Sahu, “[t]wo sets of records⁸ were relied on to reconstruct EtO emissions to air for 1984, both prepared by UCC for the Institute and South Charleston plants as part of [the West Virginia Air Pollution Control Commission’s (‘WVAPCC’)] 1984 Emission Inventory.” [ECF No. 309-3, at 56]. But when reporting this 1984 data, UCC stressed to the regulatory agency that the figures represented on the report were “drastically overstated” and did not reflect reality. [ECF No. 309-4, at 2 (February 1986 letter from UCC’s Environmental Protection Coordinator stating that UCC was “extremely skeptical of the validity” of the information provided and “convinced that the numbers generated drastically overstate actual fugitive losses”)]. UCC has consistently argued since that these estimates are unreliable. *See* [ECF Nos. 309-1, at 31 (excerpt from UCC’s 1984 Fugitive Emissions Inventory Summary for the South Charleston plant in which UCC states that it believes that its fugitive losses were “at the worst, no more than ten per cent of the total losses reported”)]. Yet, Dr. Sahu ignores UCC’s reliability concerns and uses the estimates in preparing his model, claiming at his deposition that UCC’s explanations were “unfounded conclusionary statements” despite performing no exercises to prove the validity of his assertions. [ECF No. 310, at 17 (quoting [ECF No. 309-2, at 41])].

⁸ Dr. Sahu also used a 1984 statewide Air Facility System (‘AFS’) emissions inventory located and produced by the WVDEP. But he concedes that these reports “contain emissions inventories . . . not specific to any one facility in the state” but instead combine data from all facilities within the state into one file per annum. [ECF No. 309-3, at 56]. These reports are, therefore, not capable of reliably estimating emissions from only the South Charleston facility.

Dr. Sahu's use of patently unreliable estimates is troublesome because, as noted in *TMI Litigation*, "if the data underlying the expert's opinion [is] so unreliable that no reasonable expert could base an opinion on them, the opinion resting on that data must be excluded." 193 F.3d at 697; *see also* Langley R. Shook & James A. Tartal, *Unreliable Speculation or Sound Science Under Daubert?-Use of Expert Testimony to Model Contaminant Release and Transport*, SE55 ALI-ABA 525, 541 (2000) ("Regardless of its basis, testimony founded on critical facts that are plainly untrustworthy may be excluded because it cannot be helpful to the trier of fact."). Here, UCC has expressed concerns for almost 40 years about the validity of the 1984 reported data. Despite this, Dr. Sahu "freely admitted that he performed no validation of that 1984 data." [ECF No. 310, at 10 (citing [ECF No. 309-2, at 38 (responding to question about validation of 1984 data by stating he "made the assumption that the company would not provide misleading data to its regulator"))]; *see also* [ECF No. 309-2, at 42 (stating that he did not "do any sort of validation for the 1984 data")]. But the mere fact that UCC reported the data does not permit Dr. Sahu to use it in estimating exposures without any scientific underpinnings to support his decision, especially in light of UCC's reliability concerns at the time of submitting the reports. *See Colon v. BIC USA, Inc.*, 199 F. Supp. 2d 53, 79 n.18 (S.D.N.Y. 2001) ("Some courts have stated that an opinion based solely on the defendant's data is not reliable.") (citing *Pestel v. Vermeer Mfg. Co.*, 64 F.3d 382, 384 (8th Cir. 1995) (refusing to allow expert to rely on testing done by manufacturer because he had not developed, participated in, nor supervised the testing)); *see also*

Clay v. Ford Motor Co., 215 F.3d 663, 676 (6th Cir. 2000) (“While there is a certain logical appeal to the notion that [an expert’s] opinion must be reliable if it rests upon data produced by the defendant, the notion does not withstand close consideration.”) (Ryan, C.J., dissenting); *Dow AgroSciences LLC v. Nat’l Marine Fisheries Serv.*, 707 F.3d 462, 473 (4th Cir. 2013) (stating that the proponent of data “need not have analyzed or explicitly rejected every data source presented, but it surely was required to ‘cogently explain why it ha[d] exercised its discretion’ in relying on a data set that was so highly criticized”) (internal citation omitted).

In addition to failing to take any steps to validate the estimates from the 1984 reports, Dr. Sahu concludes—“without citing to any counter-vailing site-specific data”—that UCC had *underestimated* the fugitive emissions estimates provided to the WVAPCC. *Id.* at 10 (citing [ECF No. 309-2, at 23 (replying “no” when asked whether there was any site-specific information he had relied on for his conclusion that UCC had underestimated its fugitive emissions numbers)]); *see* [ECF No. 309-2, at 24]. Sahu’s unsupported speculation as to the validity of UCC’s reported figures does not lend to his opinions’ admissibility. *See* Fed. R. Evid. 702 advisory committee’s note (2000 amends.) (“The trial judge . . . must find that [the proffered testimony] is properly grounded, well-reasoned, and *not speculative* before it can be admitted.”) (emphasis added); *see also* Itzhak E. Kornfeld, *A Postscript on Groundwater Modelling: Daubert, “Good Grounds,” and the Central Role of Cross-Examination*, 29 Tort & Ins. L.J. 646, 646 (1994) (“When the model is speculative, show it for what it is. Defective.”).

Ultimately, Dr. Sahu’s modeled EtO concentrations for the year 1984 are “over *100 times greater*” than the estimates for some of the other years in the potential class period. [ECF No. 310, at 16 (emphasis in original)]; *see* [ECF No. 390-3, at 78–79 (Tables 10–12)]. This prodigious inconsistency should prompt most experts to re-evaluate the validity of their underlying data. Dr. Sahu, however, stands by his position in his rebuttal report, claiming that because Defendants failed “to refine and create their own and better inventory,” his use of “reported emissions, along with engineering [judgment] when no data was available and where gaps had to be filled” renders his analysis “reliable, robust, and a likely under-prediction of actual emissions.” [ECF No. 309-6, at 15]. But this statement is contrary to logic, for an estimate cannot be both “a likely *under-prediction*” and a *reliable* estimate of total exposure based on those emissions.

b. 1985–1989

Next, for the period of 1985 through 1989, Dr. Sahu relies on the Toxic Air Pollutant Facility Registration Summary Sheets for both Facilities, which he “interpret[s] as emissions for year 1989.” *See* [ECF No. 390-3, at 58]. He then assumes that emissions for 1985 through 1988 “were the same as 1989.” *Id.* But this approach also drastically affects the reliability of Dr. Sahu’s estimates because the reported data on which Dr. Sahu relies did not reflect reality.

The summary sheets used by Dr. Sahu clearly indicate that the values listed are “estimates for maximum toxic air pollutant emissions,” not *actual* emissions. *See* [ECF No. 309-1, at 27 (Figure 5)]; *id.* at 26 (“These Registration Summary Sheets

reported *potential*, or maximum possible emissions of these facilities to WVDEP, not actual emissions.”) (emphasis in original). This was also echoed in a July 1990 office memorandum from an operator at the Institute facility. [ECF No. 309-8, at 2 (describing the emissions to be reported on the registration sheets as “*maximum* emissions based on *unit design capacity*, not on *actual* production rates and hours”) (emphasis in original)]. When confronted with this information, Dr. Sahu stated that it was his understanding that the figures represented on the summary sheets “were the estimates for actual emissions.” [ECF No. 309-2, at 35]. He further stated that he had reached this conclusion solely because the values listed on the summary sheets changed over time, and in his opinion, if those values were “based on maximum production capacity,” then they “would not change over time.” *Id.*

Dr. Sahu addresses the issue of data stagnation in his rebuttal report, stating that he had “been able to locate little data in the historical record which [spoke] to the mass of EtO emissions from either facility from 1985–1988.” [ECF No. 309-6, at 28]. He argues, however, that “EtO emissions were not zero in this period” and that he fills the data gap “consistent with standard practice” by using “data from the years proximate (in this instance, 1989).” *Id.* He further claims that “[t]here were no process changes [he] could identify in the record between 1989 and the gap years,” which would make his assumption reasonable. *Id.* However, he makes no mention of the true issue here: that the data he uses for these years was based on the potential maximum emission rates, not actual emissions.

Dr. Sahu’s reliance on these “worst-case emissions rates,” [ECF No. 310, at 17],

damages the reliability of Dr. Sahu's report because his estimates do not "reflect site-specific data on operating scenarios" for the South Charleston facility for 1989, let alone for the entire period from 1985 to 1989. [ECF No. 309-1, at 26].

c. 1990–2019

Finally, for the 29-year period from 1990 to 2019, UCC claims that Dr. Sahu "relie[s] on emissions estimates reported by UCC pursuant to regulatory requirements and accepted all reported emissions without analysis." [ECF No. 310, at 10–11]; *see also* [ECF No. 309-1, at 27 ("From 1990–2019, Sahu relies on 'self-reported data' submitted under [Toxic Release Inventory ('TRI')] and WVDEP annual emissions inventory requirements for [UCC's] operations at the South Charleston facility.")] Dr. Sahu states in his report that "[f]or the period 1990–1995, emissions of EtO were estimated based on TRI reporting by UCC since no other sources of data were identified which documented the amount of EtO released by emission point." [ECF No. 309-3, at 62]. The emissions for "each individual fugitive and point source" for this period "were assumed to be in the same proportion as 1996." *Id.*

In his responsive report, Machado argues that these regulatory reports "were not intended or designed for exposure reconstruction," and Dr. Sahu's use of them for that purpose presents serious methodological issues. [ECF No. 309-1, at 8].

Regulatory submittals are not always specific as to the source of emissions. For example, TRI reports "stack" and "fugitive" emissions but does not specify which stacks or which fugitive emissions. Sahu attempts to allocate emissions from regulatory submittals to various sources. For example, he states, "*Emissions from 1990–1995 for each individual fugitive and point source were assumed to be in the same proportions as 1996.*" . . . This would only be true if each emission source was used in the same proportion to each other each year, which is not

the case since different sources are related to different product manufacturing, and the amount of each type of product produced depends on customer demand. . . . accurate year-by-year estimates are critical, and Sahu’s model results that use average emission allocation factors rather than year-specific allocation factors misconstrue actual exposure. In conclusion, it is evident that Sahu’s wholesale adoption of regulatory submittals without any rigorous review of intent and purpose resulted in the improper application of emission inventories for purposes of exposure reconstruction.

Id. at 34. (emphasis in original).

In response to this criticism, Dr. Sahu attacks Machado’s findings, claiming that Machado “attempts to clumsily discredit the emissions that his own client reported” and used particular evidence to make “purposefully misleading” arguments. [ECF No. 309-6, at 14]. According to Dr. Sahu, “all [Machado] presents is an illogical phobia of anything that may have some contact with regulatory agencies, and an impossible standard that plaintiffs in litigation must somehow go back in time and force companies to collect better contemporaneous data.” *Id.* He claims that Machado is “off base when he claims that inventories reported to agencies have no use in estimating exposures,” because “WVDEP believes [required annual emissions] inventories are reliable for modeling toxic exposures and eliminating risks.” *Id.* at 15–16 (internal citation omitted).

Overall, this issue comes down to a lack of validation. “Generally, scientific methodology involves ‘generating hypotheses and testing them to see if they can be falsified.’” *Nease*, 848 F.3d at 232 (quoting *Daubert*, 509 U.S. at 593). As noted in *Coleman*, “the accuracy of [a] model depends upon the rigor applied in the input gathering process.” 2013 WL 5461855, at *24. It is troubling that Dr. Sahu has

repeatedly failed to make any attempt to validate the available data or the assumptions he makes in determining his estimated exposure, despite agreeing that it is “an important part” of his role in this case to determine whether the data presented “is, in fact, valid and appropriate” for its used purpose. [ECF No. 309-2, at 97]. Not only that, but Dr. Sahu has also repeatedly failed to validate the data he used in *forming* his hypotheses. *See id.* at 74–75 (Dr. Sahu responding that he “stuck pretty much to reported information” in response to being asked whether he “under[took] any sort of validation process with any of the emissions data that ended up as inputs” in his model); *id.* at 150–51 (confirming that he did not do “any kind of verification or validation of the 1984 numbers” because he “trusted the emission reporting by UCC in 1984”); *id.* at 160 (again confirming he did no “sort of validation for the 1984 data”); *id.* at 201 (admitting that he did not test the validity of his conclusion that nationwide NATA concentrations were “more conservative” than statewide concentrations). And while trained experts can, and often do, extrapolate from existing data, the court is not required to automatically treat such data as reliable or accurate. *See Gen. Elec. Co.*, 522 U.S. at 146. As such, it is nondeterminative that Dr. Sahu appeared to have used data from UCC’s emissions estimates because he failed to validate it. The use of this unvalidated input data—which Dr. Sahu assumes remained unchanged throughout multi-year periods—evidences that there are serious flaws not only in his methodology but also in the conclusions that he has reached. *See id.* (stating that “conclusions and methodology are not entirely distinct from one another”). This continued failure to validate, again,

casts doubt on the reliability of Dr. Sahu's model.

2. Emission Dispersion Characteristics/Meteorological Data

Next, the court has grave concerns about the meteorological data that Dr. Sahu relies on in determining emissions dispersion characteristics.

Once a substance is released into the atmosphere, downwind dispersion of it is largely driven by meteorological considerations. In reality, dispersion of the material takes place on both a very small (micro) scale and on a larger (macro) scale. On the microscale, small areas of turbulence, referred to as eddies, are generated by nearby structures (trees, buildings, etc.), thermals, and split-second variations in wind speed and direction. . . . A majority of air dispersion models predict downwind concentrations using mathematical algorithms that describe macroscale effects. Wind speed, wind direction, ambient temperature, and stability are the most-common inputs to these simplified Gaussian models. Wind speed drives the rate of transport, degree of shearing, and instantaneous concentrations. Wind direction determines which geographical areas are exposed; ambient temperature influences the rise of the material above the ground; stability dictates the level of mixing of the material with the ambient air.

Wilson et al., *Tracking Spills and Releases*, *supra*, at 390.

According to Dr. Sahu's report, "[f]or modeling of emissions from the South Charleston facility source units for all years between 1984 and 2019, hourly on-site wind speed and wind direction data collected between December 5, 1985[,] and December 5, 1986[,] were combined with twice-daily upper air radiosonde data⁹ from

⁹ According to the National Oceanic & Atmospheric Administration, a radiosonde is "a small instrument package that is suspended below [a] balloon filled with either hydrogen or helium," and "[a]s the radiosonde is carried aloft, it measures pressure, temperature, and relative humidity." *Radiosondes*, Nat'l Oceanic & Atmospheric Admin., <https://www.noaa.gov/jetstream/upperair/radiosondes> (last visited Mar. 12, 2024). Additionally, "[b]y tracking the position of the radiosonde in flight via GPS (Global Positioning System), measurements of wind speed and direction aloft is also obtained." *Id.*

Huntington, [West Virginia]¹⁰ to create a single annual meteorological dataset for input.” [ECF No. 309-3, at 76]. For the 39 non-flare Institute sources, Dr. Sahu models emissions for all years between 1984 and 2019 by using “a similar annual dataset” that “was created using hourly on-site wind data collected between December 6, 1985[,] and December 6, 1986[,] at a site adjacent to the Institute facility,” which was also combined with the same upper-air radiosonde data from Huntington. *Id.* “Hourly wind data from Yeager Airport (located 7 km east of the [South] Charleston facility) were used for hours in which the on-site data for either [South] Charleston or Institute were missing and for other required meteorological parameters such as atmospheric pressure, relative humidity, cloud cover, and precipitation.” *Id.* at 76 n.85. Surface characteristics, also referred to by Dr. Sahu as “land-use” characteristics, “were developed for the areas surrounding each of the three locations where hourly wind data were collected (South Charleston, Institute, and Yeager Airport).” *Id.* Additionally, for “the two elevated Institute flares,” Dr. Sahu collected hourly meteorological data from the Yeager Airport “for each year between 1990 and 2019,” which was then “combined with twice-daily upper air radiosonde data collected at Huntington, [West Virginia] for 1990 through 1994, at Dulles Airport, [Virginia]¹¹ for 1995, and at Roanoke, [Virginia]¹² for 1996 through 2019.” *Id.* at 77. This information was then used to model emissions for these flares for 1984 through 2019.

¹⁰ The straight-line distance between Huntington and South Charleston is approximately 41 miles, while the straight-line distance between Huntington and Institute is approximately 37 miles.

¹¹ The straight-line distance from Dulles Airport, Va. to Institute, W. Va. is approximately 239 miles.

¹² The straight-line distance from Roanoke, Va. to Institute, W. Va. is approximately 126 miles.

“Surface (land-use) characteristics . . . were developed for the area surrounding the Yeager Airport met tower,” and because “[c]omplete hourly surface meteorological data were not available at the Yeager Airport for 1984 through 1989,” Dr. Sahu uses meteorological data from 1990 for those years. *Id.*

The Defendants argue that Dr. Sahu’s methodology is flawed because he “applie[s] three different, contradictory sets of meteorological data taken from different locations in a single year (1986) to model emissions and add[s] the results together.” *Id.*; *see also* [ECF No. 328, at 16 (stating that “the model Dr. Sahu elected to use, AERMOD, is only capable of modeling one set of meteorological data at a time. . . . So, Dr. Sahu ran a separate model for each set of meteorological data he chose to use[] and added them on top of each other.”)]. “Specifically, he use[s] data from outside Kanawha Valley to model flare emissions and data from inside the valley to model non-flare emissions and sum[s] them together—assuming that the non-flare and flare meteorological data sets impacted all modeled locations across the modeling domain equally, even if the modeled receptors differ in terms of whether they are inside or outside of the valley.” *Id.* (citing [ECF No. 309-3, at 76–77]). Additionally, Defendants assert that Dr. Sahu has done “nothing to ensure that the meteorological data was consistent.” [ECF No. 328, at 16]. I agree.

Dr. Sahu’s method was to use three different sets of meteorological data from surrounding geographic locations, which is ultimately not representative of the South Charleston facility. The EPA promulgated guidelines for its own air quality modeling, which can help provide guidance in this case. It provides that “meteorological data

used as input to a dispersion model should be selected on the basis of spatial and climatological (temporal) representativeness.” 40 C.F.R. Pt. 51, App. W § 8.4(b); *see also Marias v. EPA*, 906 F.3d 1069, 1077 (D.C. Cir. 2018) (explaining how the EPA’s air quality modeling guidelines assess representativeness in response to a challenge to the EPA’s designation of an area as having high levels of sulfur dioxide) *order clarified*, No. 16-1314, 2019 WL 1096761 (D.C. Cir. Feb. 22, 2019). The EPA looks at numerous factors to determine whether the input data is representative; these factors include “(1) the proximity of the meteorological monitoring site to the area under consideration; (2) the complexity of the terrain; (3) the exposure of the meteorological monitoring site; and (4) the period of time during which data are collected.” *Marias*, 906 F.3d at 1077 (*quoting* 40 C.F.R. pt. 51, App. § 8.4(b)).

Applying this reasoning to this case, I do not find that the meteorological data Dr. Sahu uses when determining emissions dispersion characteristics is representative of the South Charleston facility, rendering it unreliable. First, the three different locations have various wind patterns and directions, which were then combined to create cumulative exposure levels. Wind directions are highly variable; however, a single location—such as Plaintiff’s residence—cannot have multiple wind speeds coming from various directions at one specific time. This in turn, renders the methods Dr. Sahu uses as unreliable because the model results in elevated cumulative exposure levels. This approach regarding meteorological data does not “tie with reality.” *Coleman*, 2013 WL 5461855, at *26.

This issue is compounded when considering the geographical terrain of the

area. West Virginia is a mountainous state, and the South Charleston facility is located in the Kanawha Valley. Ignoring the topographical characteristics of the region and using data from outside the valley without considering such terrain further indicates flaws in Dr. Sahu's use of such meteorological data as input to his dispersion model.

Finally, Dr. Sahu failed to assess year-to-year variability in wind speed or wind direction within the Valley, as he consistently applies the meteorological data for South Charleston from 1985–1986 to all of 1984 through 2019. As such, the period of time upon which his data was collected presumes that there have been no changes to the meteorological conditions near the South Charleston facility in the past 33 years. Ultimately, it fails to accurately explain or predict the concentrations, and it leads to further uncertainty in his models.

Dr. Sahu responds to the criticism of his model by arguing that “[i]t is entirely appropriate to use two different met data sets to disperse emissions from two sources when using a single met data set would be inappropriate to disperse emissions from one of the sources.” [ECF No. 309-6, at 37]. He claims that “[t]here is nothing inherently incorrect about adding the contributions from all the modeled sources to determine the total impact” and paints Machado's dissenting opinions as “speculative and unsupported rhetoric.” *Id.* And while he concedes that Kanawha Valley “may have some differences” in wind direction and speed at different heights—“especially close to the ground”—he argues that “the data from ‘outside’ the valley will typically be a reasonable representation of the flow in the wider area.” *Id.* Dr. Sahu also argues

that “[i]t is highly unlikely that modeled annual averages will vary much using ‘other’ data, even if such additional on-site data was available.” *Id.* at 36.

I find Dr. Sahu’s response unpersuasive, especially in light of his failure to perform any additional testing to rebut or otherwise respond to Machado’s assertions. One cannot reliably determine that potential variances would be minimal without having a legitimate basis for that determination. Ultimately, it does nothing to quell the court’s concerns as to the reliability of Dr. Sahu’s model and opinions.

3. Background EtO Levels

Finally, the Defendants argue that Dr. Sahu lacks a reliable method to estimate background levels¹³ of EtO in the community. *See* [ECF No. 328, at 17 (“Dr. Sahu’s opinions concerning background levels of [EtO] are . . . unreliable because he selected a surrogate figure for background that was developed by the [EPA] that does not even attempt to represent background, and Dr. Sahu ignored numerous, relevant alternative sources.”)].

In his report, Dr. Sahu admits that although he is “aware of some recent ambient monitoring at a few locations designated as background,” he feels that “the measured concentrations at these locations do not represent background” because (1) “the designated locations were selected with no underlying analysis or basis for their suitability to represent background concentrations,” and (2) “even at these locations[,] monitoring has not been done for an extended time period[,] and therefore[,] annual

¹³ “Background” refers to exposure levels experienced by the general population. *See Rhodes v. E.I. du Pont de Nemours and Co.*, 253 F.R.D. 365, 374–75 (S.D. W. Va. 2008) (Goodwin, J.).

average monitoring data are simply not available.” [ECF No. 309-3, at 80]. However, he fails to indicate the source of this information or offer any additional details that would invalidate such data. Instead, he states that in determining a background level, he “relied on EPA’s NATA database to obtain annual average concentrations of EtO,” which showed an “annual average concentration for all U.S. exposures” of 0.000216 ug/m³. *Id.* (citation omitted). He then compared that figure “against predicted model concentrations” to determine that “the class representatives were chronically exposed to levels far higher than national background levels.” *Id.* “In fact,” he claims, “in all years of their residency except 2019, the Class Representatives were exposed to levels that were more than 100x higher [than] national annual average background exposures.” *Id.*

UCC argues that despite the EPA’s statements that they “don’t know what background levels for EtO are” and “can’t put an exact number on it,” Dr. Sahu “used one cherrypicked figure from the EPA to approximate the background concentration of EtO.” *Id.* at 18–19 (citing [ECF No. 309-12, at 3]). According to UCC, this “surrogate figure” fails to “consider ‘ethylene oxide in the outdoor air that is not clearly linked to a particular industrial facility,’” and cannot, then, be considered a background level. *Id.* at 19 (citing [ECF No. 309-12, at 2 (defining the term “background ethylene oxide” as EtO “in the outdoor air that is not clearly linked to a particular industrial facility”)]). UCC also claims that Dr. Sahu’s figure “has nothing to do with the proposed class area” because Dr. Sahu used a national annual average instead of one specifically linked to the South Charleston area. *Id.* (citing *Dodge v. Cotter Corp.*, 328

F.3d 1212, 1225–26 (10th Cir. 2003) (stating that background concentration measurements should be “tied” to the proposed class area)).

Next, UCC disagrees with Dr. Sahu’s contention that the locations designated as background in recent ambient air monitoring were “selected with no underlying analysis or basis for their suitability to represent background concentrations.” [ECF No. 309-1, at 47 (citing [ECF No. 309-3, at 80])]. Instead, Machado asserts that the WVDEP, who completed that monitoring, “clearly explains” in their monitoring report that “[t]he background sites were chosen in areas where there were no known sources of EtO.” *Id.* (citing *Ethylene Oxide (EtO)*, W. Va. Dept. of Env’t Prot., <https://dep.wv.gov/key-issues/Pages/EtO.aspx> (last visited Mar. 18, 2024)). The WVDEP’s background measurements “ranged from 0.06 ug/m³ to 1.71 ug/m³, with several of the background samples indicating concentrations higher than those taken onsite or around the [F]acilities.” *Id.* In fact, “[t]he highest measured background concentration, 1.71 ug/m³, is higher than a *majority* of [Dr.] Sahu’s modeled concentrations at the class representative location.” *Id.* While Machado states that he “recognize[s] that WVDEP monitoring consists of snapshots in time, not an annual average, and concentrations can fluctuate in a year,” he proposes that WVDEP’s data suggests that background EtO concentrations not linked to a particular facility “cannot be ignored.” *Id.*

Machado also takes issue with Dr. Sahu’s decision to use NATA annual average concentrations for all U.S. exposures as his EtO background level. *Id.* at 48. He argues that “NATA provides screening estimates of concentrations from modeling

of emission sources and does not incorporate measured concentrations of background EtO in model results.” *Id.* Further, “NATA modeling extends to the far reaches of the continental US, in areas remote from anthropogenic sources of EtO rather than the populated area represented in [Dr.] Sahu’s modeled domain.” *Id.*

In response to these criticisms, Dr. Sahu argues claims that “[b]ackground, *as relevant for this matter*, is the concentration of ethylene oxide at the locations of interest, *but for* the contributions from Defendants’ plants and activities.” [ECF No. 309-6, at 17]. “Background is not, as Mr. Machado (and others) presume, some measurement (or questionable quality) conducted by an agency for some other unknown purpose.” *Id.* at 17–18. Because “no agency[—]or anyone else for that matter[—]has collected continuous measurements for one year or more in any location relevant to Plaintiffs’ background,” Dr. Sahu contends that his use of the NATA data is “appropriate and common.” *Id.* at 18. He claims that he “looked at AirToxScreen results for the nationwide average across the entire US, consistent with its purpose” and not “to assess what any Plaintiff or Class Members’ exposures are.” *Id.* Additionally, Dr. Sahu argues that “[l]ike his other non-material and trivial criticisms,” Machado’s assertion that Dr. Sahu’s use of a single year of NATA data in determining background concentration levels for additional years is indicative of Machado’s misunderstanding of “the relevant issue.” *Id.* at 18–19. He claims that he is “not aware of any EPA guidance that contradicts” his approach. *Id.* at 19. I disagree.

The EPA has expressly indicated important limitations to user applications of its NATA data sets:

EPA suggests that the results of this assessment be used cautiously, as *the overall quality and uncertainties of the assessment will vary from location to location as well as from pollutant to pollutant. In many cases more localized assessments, including monitoring and modeling, may be needed to better characterize local-level risk. . . .*

The following are important specific limitations to recognize:

- The results apply to geographic areas, *not specific locations. . . .*
- The results apply to groups, *not to specific individuals.*
- The *results are restricted to the year of the analysis* since emissions for that year were used. . . .
- The *results do not fully reflect variation in background ambient air concentrations.*
- The *results might systematically underestimate ambient air concentration for some compounds.*

National Air Toxics Assessment: NATA Limitations, U.S. Env't Prot. Agency, https://19january2017snapshot.epa.gov/national-air-toxics-assessment/nata-limitations_.html (last visited Mar. 12, 2024) (emphasis added) [hereinafter *NATA Limitations*]. Moreover, “courts have consistently excluded expert testimony that ‘cherry-picks’ relevant data.” *Lipitor (Atorvastatin Calcium) Mktg. v. Pfizer, Inc.*, 892 F.3d 624, 634 (4th Cir. 2018) (quoting *EEOC v. Freeman*, 778 F.3d 463, 469 (4th Cir. 2015) (collecting cases)). This is because expert testimony based on cherry-picked data “does not reflect scientific knowledge, is not derived by the scientific method, and is not ‘good science.’” *Id.* (quoting *In re Bextra & Celebrex Mktg. Sales Practices & Prod. Liab. Litig.*, 524 F. Supp. 2d 1166, 1176 (N.D. Cal. 2007)).

In conflict with the EPA’s statements that “the results are restricted to the year of the analysis,” *see NATA Limitations, supra*, Dr. Sahu cites to only the 2019 AirToxScreen results and seems to compare these 2019 results to each year from 1984 to 2019. *See* [ECF No. 309-3, at 80]. Moreover, it is clear from the EPA’s numerous

statements that the background concentration levels of EtO are not known with any level of certainty. And despite there being a local monitoring report replete with information that Dr. Sahu could have potentially used in calculating a reasonable background level, Dr. Sahu chooses instead to ignore local data in lieu of using a national average from a single year—the validity of which is questioned even by the very entity that compiled the data. Although Dr. Sahu believes that the nationwide estimate is much more “robust and representative,” he provides no evidence to back up that statement aside from asserting that the “locations designated as background were not selected using any criteria designed to ensure that they are robust or representative estimates.” Such self-serving statements do little to alleviate the court’s concerns about his methodology.

“In the absence of individualized personal monitoring data, toxic tort litigants [attempt] to use air-modeling techniques as evidence of a plaintiff’s actual exposure.” Jose A. Berlanga & Nancy J. Brown, *Establishing Tort Liability with Regulatory Tools? The Utility of Air Modeling as a Surrogate for Monitoring Data*, 3 Env’t & Energy L. & Pol’y J. 1, 27 (2008). However, “while models are most reliable for estimating the highest concentration occurring within an area, they lack the fundamental capacity to show actual exposure to an individual at a precise location.” *Id.* at 29. Further, “the sheer complexity of [the] mathematical computations [involved] suggests that those looking to air modeling data should proceed with caution, especially when applying models to ambient individual exposures.” *Id.* at 30.

In resolving challenges to models, most courts perpetuate the pervasive misunderstanding and assume that since the model is mathematical, it is correct. As a result, they pass the model through the system without much, if any, scrutiny. The dominant trend in the courts follows this deferential, deterministic path, even in cases when there are reasons to suspect that the model may have significant problems.

Wagner et al., *Misunderstanding Models*, *supra*, at 320; *see also* Swinehart, *Remedying Daubert's Inadequacy*, *supra*, at 1296 (arguing that models “are not precise, and their remarkable value should not obscure the judicial inquiry into model reliability.”). I decline to take that path and find instead that where, as here, a plaintiff's claim rests in large part on one expert's determination of exposure, “[a]n in-depth data investigation, a searching historical analysis, an excruciating attention to detail, and a methodology designed to wring error out of the process seems especially *apropos*.” *Coleman*, 2013 WL 5461855, at *24. Despite these methodological principles, Dr. Sahu has presented us with a model that is “a speculative conglomeration of data that is unreliable on the question of exposure” surrounding the South Charleston facility. *See id.* at *25. “This basic methodological flaw infects his entire analysis” and “makes for a patently unreliable measure of significant exposure to harmful substances, which is the central issue” in this case. *Id.*

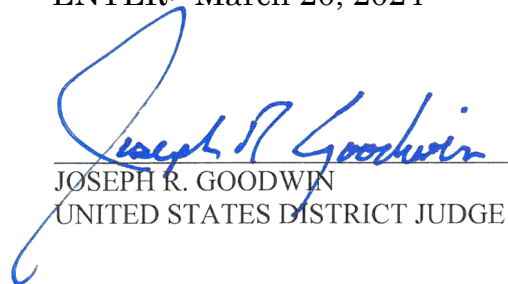
I find that the opinions of Dr. Sahu are not based upon sufficient facts or data. However, even if that underlying data were sufficient, Dr. Sahu *application* of the data to the facts of this case was methodologically flawed, full of unsubstantiated assumptions, and not scientifically sound. Therefore, I find that his opinions are not the product of reliable principles and methods, and these methods have not been

reliably applied to the facts of this case. As such, his opinions should be excluded for a lack of reliability under *Daubert*.

IV. Conclusion

Having considered the entirety of the challenge to Dr. Sahu, I **GRANT** the Defendants' Motions to Exclude his testimony and opinions. [ECF Nos. 309, 327]. The court **DIRECTS** the Clerk to send a copy of this Memorandum Opinion and Order to counsel of record and any unrepresented party.

ENTER: March 20, 2024



JOSEPH R. GOODWIN
UNITED STATES DISTRICT JUDGE